



UNIVERSITÉ LUMIÈRE LYON 2

Theme:

Expérimentation et évaluation d'outils d'OCR et d'OLR.

Présenté par Edina ADJARO PATOUSSI 30 Aout 2024

Tuteur laboratoire:

Tuteur enseignant:

Julien VELCIN Ludovic MONCLA

Valentin Lachand-Pascal







sommaire

OU OCR

CONTEXTE ET PROBLÉMATIQUE EXPÉRIMENTATION DES OUTILS. PRÉSENTATION DE LA PROPOSITION DE MÉTHODOLOGIE STRUCTURE D'ACCUEIL **OBJECTIF** RÉSULTATS OBTENUS ÉVALUATION DES RÉSULTATS SEGMENTATION OU OLR RECONNAISSANCE DE CARACTÈRE **CONCLUSION ET PERSPECTIVES**





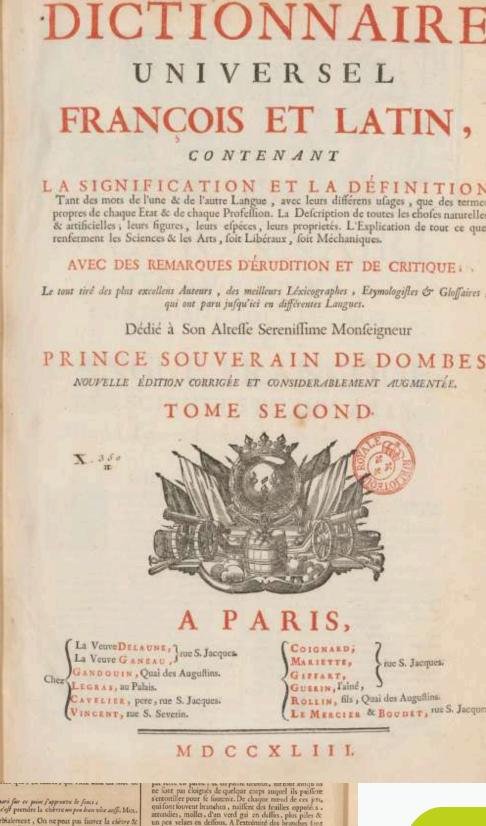
Contexte et problématique.

Le **Trévoux** est un dictionnaire français datant du 19ème siècle, qui existe en plusieurs versions.

Étant une œuvre importante, il a suscité l'intérêt de **linguistes** et a été l'objet de diverses études. Cependant, ne disposant pas d'un format numérique (.txt) il est difficile d'en faire une analyse approfondie.

C'est dans cette optique que ce stage a été initié, afin d'expérimenter des outils d'OCR et d'OLR sur le Trévoux

Comment surmonter les difficultés liées à l'**OCRisation** du Trévoux tout en respectant la structure du document ?









Présentation de la structure d'accueil



Deux équipes :

- Équipe **DMD** :
 - Intelligence artificielle
 - Machine learning
 - Data mining
- Équipe **SID** :
 - Big data management
 - Big data analytics
 - Analyse en ligne (OLAP)
 - Sécurité des données



Douze équipes de recherche structurées en six pôles de compétences:

- Données, Système et Sécurité
- Informatique Graphique et Géométrie
- Images, Vision et Apprentissage (équipe IMAGINE)
- Interactions et cognition
- Algorithmique et Combinatoire
- Simulation et Sciences du Vivant







- Tester et comparer les outils d'OLR/OCR sur les pages de **Trevoux** de 1704 et 1743
- Constituer un jeu de données annoté pour l'entraînement de modèles d'OLR/OCR avec Label Studio ou Prodigy si nécessaire.
- Développer une méthode permettant de comparer la sortie de ces outils avec la vérité terrain.

Trevoux 1704

A CE

A CUSATEUR, Accusarato s. fubft. | Accusa a, fignific aufli fimplement, Declarer. nafe, & fem. Celui ou celle qui acculé, ou qui poerint quelcus en Julice et imisellement. A.-salette. Acculatris. Par le Drote civil in 's avoir point d'acculatrese public. Chaque parti-culer, foit qu'il suit interet au crime public, ou culer, feit qu'il uit interèt au crime public, ou non, pouvoir acculer, de conclure au châriment de l'acculé. En France il n'y a que le Procureur General, ou fes Subbitunts prépoiez dans chaque siege, qui le puillent confituer aesufateurs; c'est à eux feuls à qui appartient la vengeance publique. La partie civile ne peut conclure qu'à la réparation, de aux interès, de non pas à la pu-nition du criminel. En quelque lieu que fe trou-me Partielle, il renouver un desfateur.

ACC.

Joge, & un Boarreau. L. x M. A. 1.x. Cette fem-me cit une dangereule accufarrice. Son accufarri-er parus fort animbe courte lui. Au dernier jour nos pechez se presenteront comune autant de cruels accufateurs. N. 1.c. o. 1. A. c. C. U. S. A. T. I. F. I. m. Terme de Geammaire. Cett le quatrième cas des noms qui se déclinent. Accufands casur, Accufatours. Il marque & designe le sujet où passe l'action du verbe, ou de la préposition. Un verbe actif regit l'accufasss. Il y a des prépositions qui demandent après elles un accufatif. En François l'accufasss et sim-biable su nominatif.

CUSATION. C. f. Delation en Juftice de prévenir une accujation li redoutable; ou de la repouller vigourealement, fi elle est déja formère. An La Ne. Sufciter une accujation ca-pitale. Il y a vingt chefs d'accujation contre ce criminel. L'accujation des crimes privez n'è-tost rocevable par le Drois R, qu'en la bouche de ceux qui y avoient interèt : pour les crimes publics l'accujation pouvoit être intendée par quicouque la vouloit entreprendre. Il liguifie aussi, Confession. Confessio. Il faut fai-re une sincere accujation de nos pechez devant Dieu.

Accusation, le dir aufil des legères faures dans les complimens ordinaires. L'accusation que vous me faires de n'avoir point fongé à vous en vêtre absence, est mai fondée.

Accus E.R. v. act. Intenter une action etiminelle courte quelqu'un, foir en son son, joir sous le nom de la partie publique qui est toujous le Procureur General, ou son Substitut.

Accuser, Il n'appartiere qu'un aux d'accuser. Accujare. Il n'appartient qu'au mari d'accujer fa femme d'aduliere. On a accujé de concultion un tel Officier. Tout mari prudere doit diffi-muler les infidelires de son évoute : parce qu'on

Enunciare, Lapoure, Il a accufe 50, de point au piequet i il a accufe la récepcion de malettre, pour dire ; il a dit qu'il avoit 50, de point; qu'il avoit reçû ma lettre.

qu'il avoit requima tetre.
Ac custa, avec le prouom perfornel, Se dé-clarer coupable. Infirmalare fe. Ce etiminel s'en accufe lui-même. Il faur qu'un penitent s'accu-fe franchement de fes pechez à la Coofulfon. Les perfecueurs femblent s'accufer de n'ene pas hien convaincus eux-memes de la force, & de l'evidence de leurs railòus, pailqu'ils em-

pioyent la violence. C L.

A C C U S É, le prend quelquefois fubliantivement.

L'arcafé a donné de bors reprodres contre les témoiss. On doit entendre l'arcafé, à peine de mullité du jugement. L'accafé ne peur point réfigner, quand le crime emporte la privation de l'innuilition l'on constrair l'arcafé con l'archée. B o u c u. Par les dures loix de l'Innuilition l'on constrair l'arcafé à c'erre. de l'Inquission s'on concraint l'areas, à s'accu fet lui-même du crime qu'on lui suppose. I si o DE GOA. L'acrufé n'est point reçà à accus ton accusaceur, ni à user de recrimination, avan qu'il se lois purgé. De Laos,

tel Seignout. Ce mot elt un compose de census inguisant tente annuelle, ou cens. A CE N S LM E N T. f. th. L'action d'acenser. Conduitto, locatio fundi. L'acensement d'un

Fundum locare. Un Seigneur feodal nemfe une nement pour la débonche. terre à une telle quantité de cens, ou de redevan. CANSER, figuifie aufli en plufieurs Provinces

Donner à ferme moyennant un currain prix & redevance annuelle pour un certain temps. Il u'a pû acenfer cette métairle, elle lui eft demeutee fur les bras faute de fetunier. Ce mot vient du

CERBE. Adj. Ce qui eft kere, verd, apre. Acer bus. Les Medecins tiennem que ce gout ell mi-toyen entre l'aigre, l'acide, èt l'amer. Its appelleux duvin acerbe, da vin fair de railins qui ne ione pas encore meurs. Tous les fruits avant leur mann-rité, ont un goût acerbe. La faveux acerbe, est goût . l'appetit de la chair. Canes, accipitres

A C H.

ficurs animaux. Le fenriment le plus probable eft de ceux qui ditent que les aceratautes font ces glandes quis élevent dans la mantice des brebis de des chevres , lorsfoy élles four pleines , ét qui font ainfi appellées , parcequ'elles font faltes en forme de coupe ou de godet. Ce qu'on ne remarque pes dans les femelles des autres azémaux, non plus que dans la femmle.

ACLT ABULE, figuific encore une cettaine me-fure dont les Apoutaires fe fervou pour les cho-fes liquides. Voyes Cocyle, Cosyledon. ACETABULUM, C.m. Sotte de plante,

appelie autremert ambilions Veneris. Il y en a de a. fortes: l'un dont les feuilles font exeules ; & connices comme un acetabule, ou une coupe. L'autre jette une tige meur, & produit des fleurs femblables à celles de mille-pertuis. Cetre plante a les fetifiles larges, & fort épailles. Sa graine qui est un peu grotte, alts monts pro-prietez que la jouborbe.

A C H A L A N D E R. erbe actif. Attiret les ou il fe lois purgé. De Laos.

A C E.

A C E N S E. C f. Hétritage, ou ferme qu'on tient à perpetulité, ou à longues années, d'un Seigneur à certain cens & tente, ou à prix d'argant Fandaux pretis conduits. Cette métaitie est une securje d'une telle Abbaye. Il n'est pas propriétaite de cet hétritage, il le tient en accuje d'un personne c'a s'achalmater. Ou le dit suffi en badianne.

dinant, d'une personne qui a beaucoup d'intri-gues : Cotte fille est fore achalandes.

ACHALANDÉ, 2's part. paff. & adj.

emportement i attachement contratte a quesque choie. Libido a proyenfio, Il se dit plus ordinal-rement en mauvaise part. Il a un furieux achar-

de l'animonté avec laquelle on perfecue quel-du'un. Injeffarie vibennes, acrés. Ces deux Auteurs ont un farieux acharnement à la perdite mutuellement i lis fe déchirent par tout.

Tous les devots de cour sont aifez à connolère. Jamais contre un preheur ils n'ent d'achat-

Trevoux 1743

CHE

retour de cés oifcaux, c'est-à-dire, au premier printemps, ou qu'étant bonne pour la vûe, les hirondelles s'en servoient pour la rétablir à leurs petits. Sa racine est grosse comme le petit doige, fibreuse, rougeatre en dehors, & d'un jaune de fassiran en dedans. Elle jette de son collet plusieurs seuilles longues au moins de demi-pied, velues, d'un verd pale, & découpées profondément en quatre ou cinq fegmens en manière d'ailerons rangés comme dans les feuilles, com-polies & terminées par un fegment beaucoup plus large que les autres. Chaque fegment reffemble à la feuille du chêne, & il est pareillement incifé sur ses bords en ondes. Les tiges qui s'élèvent d'entre ces feuilles font grêles, divifées en quelques branches, chargées de quelques feuilles affez femblables à celles du bas, & font terminées par des bouquets de fleurs dont les pédicules communs fortent du côté opposé des branches de la tige, lesquelles branches naiffent des aisselles des feuilles, Ces fleurs sont jaunes, conleur de la teinture du faffran, à quatre pétales foutenues par un calice, à deux petites feuilles, qui tombent en même temps que la ficur s'épanouit. Des étamines sans nombre & de la même couleur des pétales entourent un pistile qui devient enfuite une filique longue de deux pouces environ fur une ligne & demie de largeur, conleur d'olive, composce de deux panneaux appliqués sur les bords d'un chas-sis à jour; c'est-à-dire, qui n'est couvert d'aucune membra-

Toute la plante donne un fac jaune & âcre, elle est très-apéritive. On en recommande l'usage à ceux qui ont des dartres, foit prife intérieurement dans des aposêmes, foit extérienrement en fomentation. On s'en fert aussi pour le maladies des yeux , mais on tempère l'âcreté de son suc avec le lait. La plante qu'on nomme perite éclaire est tout-

CHELLES, Bourg de l'îlle de France, à quatre lieues de Pa-ris, sur la Marne. Cala. Sainte Baudour, femme de Clovis II. y fonda une Abbaye de Religieuses, dans laquelle elle fe retira après la mort du Roi son mari ; & où Cloraire II. son fils fut enterré. Le Roi Robert y avoit un Palais, qu'il appelle dans un Edit Kale nostre Palatium. Du Chesne

ne, comme dans presque toutes les autres filiques. Aux deux côtés de ce chassis sont attachées des semences me-nues, longuettes, noires & luisantes dans leur maturité.

CHELONITE, L. F. Celt une pierre qui se trouve au ventre des jeunes hirondelles, qu'on estime bonne pour le mal caduc. Chelonia. Il y a une autre chelonite qui se trouve aux tormes des Indes, qui a la vertu de réfifter au venin. QuelCHE

que Juge dans son territoire est obligé à veiller à la sûresé des grands chemins.

Voit-on les loups brigands , comme nous inhumains , Pour détrouffer les loups course les grands chemins ? Bott.

Berger, dans son livre des grands chemins de l'Empire, dit que ce mot est du vieux François. La Guide des chemins enseigne les routes des grands chemins, les postes & leur distance. On dit, Paffer fon chemin, fe mettre en chemin.

Ce thot vient à caminit. Quelques-uns le font venir de capmum, qu'ils disent signifier la même chose. D'autres le tirent du Latin semira, & d'sent que chomin a été dit comme semin.

temin des arbres, dans l'Artois & dans les pays conquis en Flandres, fignific grand chemin, chemin Royal, chemin des troupes. Via militaris. Ce nom a été donné aux grands chemm par les gens de ces pays-là, parce que le Roi, en fai-fant elargir & accommoder les grands chemins, fit planter des arbres des deux côtés, d'où il revenoit deux avantages confidérables; l'un que les troupes n'avoient pas besoin de guides pour connoître les chemms les plus courts & les plus commodes, l'autre qu'on avoit toujours dans le pays une reflource pour tout l'attirail de la guerre, & sur tout de

On appelle chemin du halage, un chemin de 24 pieds de large, que les Riverains des rivières navigables sont obligés de laisset sur les bords pour le passage des chevaux qui halens ou tirent les bateaux. Ripa. On appelle chemin de traverse: Transversum iter, Tramses, me

Chemin detourne, ou qui n'est pas sur la route des grandes villes, mais qui va d'un bourg ou d'un village à un autre. ou qui n'est pas le chemes ordinaire pour aller d'un lieu à un autre, si c'est quesque grand lieu. Via deversa, devia. On appelle aussi chemin particulier, un chemin de traverse, & on l'oppose au chemin public, qui est la même chose que grand chemin.

hemin ferré,celui qui est pavé,ou dont le fond est dur & soli-de, de roche, de pierre. Via strata. Chemin de velours, celui où il y a de l'herbe, qui est sur une pelouse. Herbosa, cespititia. Chemin creux ou cavm, qui est enfoncé au-dessous du rez de chaosse. Iter depressum, pervium. Chemin fourchu, celui qui se divise pour aller en divers endroits, bivium, trevium, quatrivium, felon le plus ou le moins d'endroits où il conduit. Chemin dessieile, qui est âpre, raboteux, ou qui est mal ailé à tenir pour être coupé en plusieurs endroits; asperum, durum, difficile. Chemin bas, qui est dans la vallée, depressam. Chemin haut, qui est sur la colline. Supernum, supernum.





- Layout Parser (OLR)
- Surya (OLR/OCR)
- Laypa (OLR)
- Tesseract (OCR)

OLR + OCR </article>

<article> <vedette>CHELLES</vedette> bourg de l'i...





Segmentation ou OLR.

Definition:

La segmentation d'image est une opération de traitement d'images consistant à détecter et rassembler les pixels suivant des critères, notamment d'intensité.

Pourquoi:

- Bien Extraire la Structure des Dictionnaires
- Améliorer la Précision de l'OCR







Reconnaissance de caractère ou OCR

Definition:

La Reconnaissance Optique de Caractères (OCR) est un processus qui permet de convertir du texte imprimé en format numérique

Erreur:

■ **Délétion**: caractère manquant

word \rightarrow wrd

■ Substitution : un caractère remplacé par un autre (voire plusieurs)

word \rightarrow w0rd word \rightarrow ivord

Insertion : caractère inséré dans le texte

word -→ wordsd



Expérimentation des outils

Layout Parser

Bibliothèque Python qui permet de faire de l'analyse d'image en détectant les layouts des différents éléments d'une image en utilisant des modèles entraînés avec Detectron.

ABS.

tion. Abfalutorius. Il a une sentence absaluto es BSORBANS, Terme de Médecine. Me i

camina ad abjumendam nata. Ce foot des n'e-dicamens qui oot une puilfante vertu pour ce s-famer l'humeur amollie, & artenuée ians la d'e-foudre. On s'en fert ordinairement pour les 1-meurs, après avoir employè les émolliens & les attenuatifs. Ces médicamens font le vinaig e,

l'eau marine, la faumure, &c. S. Evn.

fumer, emporter. Absumere. Les eaux absor bene presque toute la lumière qu'elles reçoiven

du foicil. Ro H. Il est peu en utage au propre

fi ce n'est en parlant des animaux voraces : mai

il se dir au figure, & il emporre d'ordinaire u mauvais sens. Les droits de la femme one absort

tous les biens du mari. Les frais d'un técille abjer bent les plus clairs deniers de cette fuccethon

ABSOUDRE. v. act. Decharger d'une accu-

fation, de la peine d'un crime. Abjetvere, Abjet-dre un j'eutent, lui remettre les pechez dans le tribu-al de la confession. Abjendre un accuse. Dans le doute il est plus expédient d'abjendre un

On l'a ab our à pur & à plein. On die auffi dans le discours familier , en parlant d'un deffunt : que Dieu abjolue , c'est-à-dire , à qui Dieu faite

asous, ours. adj. Affranchi ou déchargé

HJDataset

ABS. ABS. Pinenes font néceffaires , p us rétablir l'efteest impertinent , in croyable , impossible 48 ABSOLUTOI & E. adj. Qui porte absolu-tion. Absolution . Il a une sentence absolutoire. ABSORBANS, Terme de Medecine. Medichofe ab urde, on en tise mille conjequer detfence ou privation de ninger de la chai abfurdes. Il prouve une choie abfurde par BSURDEMENT. adv. D'une marière il-

camina ad a fun endum nata. Ce font des mèaffaicentia carniom. L'avent la compagnée de la priere, est un des moyens les plu efficaces pour avancer notre fanchineation. Boss. L'Egli: dicamens que ont une puissante vertu pour con-famer l'hum ur a nollie, & attenuée tans la difmeurs, après avoir employé les émolliens & ordonne simplement l'abstimence le jour de lair i les attenuarif. Os médicamens font le vinaigre Mare, & non pas le jeune. Les Mécredis (or) l'eau marine la fumure, &c. S. Evr. des jours d'abjemence chez plus :urs Religieur. Les devots font aufli des abitimen es , & des ma

ABSORBE L. v. act. Engloutir, diffiper, con fumer, emprirer absumere. Les eaux absorbene presque tout la lumière qu'elles reçoiven du soieil. R. H. Il est peu en usage au propre RSTINENT. ENTE. adi. Tempérant l'égard du boire & du manger Sobrins. Le peuples du Midi four plus abjimens que ceux d il le dir au fig uté . & il emporte d'ordinaire us mauvais sens Les étoits de la semme ont absorbe

tous les biens lu m ari. Les frais d'un (celle ab/or-

ABSOU DR E. v act. Decharger d'une accu-

fation de la prine c'un crime. Abjolvere, Abjon

dre un pinite t, l'i remettre les pechez dans le

tribu: al le le coe fession. Absordre un accuse.

Dans le doute il eff plus expédient d'abjondre un

On l'a ad aus i pu & à plein. On dir aufli dans

L'abjouts fe fut a iffi par les Curez dans les Pa-

A BS TEILE f.n. Terme dogmatique. Qui

roiffes le pur le P. ques.

Morbeo fen ifiant le même.

BSTERGER, v. act. Term: de Médecin & de Chiturgie. Purger. , nette yer une playe

bene les plus clairs deniers de cette fuecession. BST BR STF, IV B. adj. Qui pringe & nettoye Smegmaticus, fmellieus. Medic ment, purga-Ce goin re a abjerbe tout ion patrimoine. La voix eft abferbie dans les voutes. A BLANC. C'est-à cire qu'e le s'y perd. La question de l'infaillit lité de Eglise absorbe toutes les au-

BSTRACT, ACTS. Terme ce Philosophi tres cont ove fes. CLAUD. Ce mot vient du Latin Il est un peu barbare en François Ce qu'on de tache par la penice de toute au te choie, an d'en avoir une connoillance fimple, & par lui meme. Species abstracta per men em ab iis reba quibus adherer. La quancité et un terme al frait, quand on la confidére en lle-même; à fans être attachée à aucun corps, quoy qu'ell ne puille lubi iter naturellement aus lui , ni li fanselle. La blancheur est un te me abstract connoillance des abfiracts on pary ent à celle de

le difeours f mili :r , en parlant d'un deffunt : que disco de , de crime Af slar ir. Ausous, le di auf i en matière civile. Un deffendeur cor :lut où surs à être renvoyé quirte & ABSOU FE. C.f. Absolution publique & fo-Amnelle jui l'doi re au peuple. Dies Abfolufaint, ou : Me cree au foir dans les Cathédrales.

l'elprit, par laquelle on confidére quelque par tie d'un tout, tans faire attention sux autres : or un détachement qui te fait par la entée de tou les accidens ou circonfrances qui peuvent ac compagner un être , pour le coclidérer mieu: en lui-mome. Aitro animi forcem aliquan abilrabentis. On confidere par ab Iraction , los que dans un mobile on confidére e mouvemen fans faire attention au corps mi. Ce font le Mathématiciens qui confiderant I. quantité fan macière, supposent dans leur emp re d'abstrac-

furde. Ab, wede. C'eft concluse abjurdeme 1;

BSURDITE, f. f. Qui contient quel 1 choie d'abiurde. Abjurde airium aut faiti » Il s'enfuivroit de grandes abjurantes d'une e l': fuppolition. La plus grande des abjurantez et la contradichon. Occille foi pana on violitei à la gens qui propolent tetieutement d'aufli gran k i

B S U S. L. m. Herbe qui eron en Egypte : 11 hauteur de quelques doigts. Ses feuilles reslet a bient à ceiles du triolet ; & les fleurs blanche ; & d'un jaune pale, produitent une temence ne ire , renfermée dans de petites cellules,

Abunthaum, lelon Malherbe ; & felon Vans r las, toù jours matculin. On le fait plus ordin i rement feminin. L'Academie Fr. le l'ait femin e Ménage veut qu'on écrive appende par un > fans doute à cause de l'étymologie. C'est e 1 plante médecinale. Il y en a de quatre fortes : l Santonique, le marin, le vulgaire ou grand Pontique, & le petit. L'absymbe comman : une tige fort branchoe , des feuilles blanches , 8 fore decouppées , co: . e l'artemilia ; ses flet r dorces & petites. Sagraine est ronde, & ditp fee comme une grappe de raifin. Sa racine () fort espatpillée. Cette absparbe vulgaire est plu en usage dans la Medecine. Pluseurs croye n que c'est la barbotine, qu'on appelle semen s'a se bien differente, Quelques-uns prétendent q e l'abjunte et l'auronne femelle. Il y a une i

alsyne, qui est temblable à la pense auronn

étant toute entailée de petite graine fort amen ,

qu'on appelle en Latin abfinibium morimum

ou feriphinen. L'absinche Santonique est la tro -

fitme elpèce , qui est semblable à l'aluine , ma s

qui est moins chargée de graine que l'autre. Cu fait du vin d'abjunthe, & de l'eau d'abjunth.

L'abfinibe est une herbe odiferante, très-amé-

PubLayNet

ABS.

fimenes font nécessaires , pour rétablir l'esto-mach affoibli par la débauene. ISTINENCE, fignific quelquefois une fimple deffince ou privation de manger de la chair abilimenten carneum. L'abilemence des viandes affaifonnée de devotion, & accompagnée de la priete, est un des moyeus les plus efficaces pour

percet, est un des moyens ses pous emacetes pour avancer noire fanctineation. Bos s. L'Eglist pedoune fimplement l'abstineation. Bos s. L'Eglist pedoune fimplement l'abstineate le jour de Laint. Mare, & non pas le jeune. Les Mercetis sont les jours d'abstineates cher plusteurs Religieux. Les devots sont aussi des abstineates, & des ma-

bent ies pass clairs demets de cette inecession.

Ce goinfre a abjorbie dans les voutes. A BLANG.
Celt-à-dire qu'elle s'y perd. La quettion de l'affilibilité de l'Eglife abjorbie outes les autres controverfes. Claud. Ce mot vient du Latin

Lettun peu barbare en François. Ce qu'on dé-

ache par la pentee de toute autre choie, afin en avoir une connoitlance simple, & par luineme. Species abstracta per mentem absis rebus usbus adharer. La quantité est un terme abrail , quand on la considére en elle-même ; &

onogifiance des abfiracts on parvient à celle des oncrets, qui est le terme opposé.

A 1 S T R A C T I O N. L L. C'est une action de

l'esprie, par laquelle on confidère quelque par-t e d'un tout, tans faire attention aux autres : ou Assous, se dit aussi en marière civile. Un deffen-In accident ou circonflances qui peuvent ac-ompagner un étre , pour le confidérer mieux in lui-meme. Atto animi fecciem aliquam bifrabentis. On confidere par abstraction, lors i ue dans un mobile on confidere le mouvement faint, ou le Mécredi au foir dans les Cathédrales L'abjoure se fait aussi par les Curen dans les Pa-

I uns faire attention au corps mû. Ce font les l'éarhématiciens qui confiderant la quantire fans satière, supposent dans leur empire d'absérac-ron des indivitibles sans parties: mais il n'est pas permis aux Physiciens de faire ces sorres d'ab-BS TEME. f.m. Terme dogmatique. Qu ne boit point de vin. Ablemins. On s'en fert dans la Théologie, pour parler de ceux qui éans la Communion se peuvent point boire de vin, à caufe de l'averson naturelle qu'ils out rastions, ni de sortir des borors de la musière. LEN. Pour bien juger d'un homme, il faut pour cette liqueur. Mr. de Meaux s'est servi de l'exemple des abstêmes , pour défendre le ret per ou pour, ou contre lui

ABU. tait oes sout neuces pat un par regime de vivre, A BOUK DE sel, maic extern. serme ne pris-cemme de vin, de falines, ecc. La dière & l'ablosophic. Ce qui choque le tera commun, qui est imperiment, i ecoyable, impossible abyardan. Preposition abyarde, quand on sipposic use chose ab, arde, o o en tire mille consequences abjurder. Il prouve une choic abjurde par une

A BSURDEMENT, adv. D'une marière ab-

les jours d'abjenence chet plusteurs Religieux.
Les devots sont aussi des abjenences, & des maétracions volontaires.

13 TINENT, and a sur a adj. Tempérant à
égard du boire & du manger. Sobress. Les
euples du Midi sone plus abjenens que ceux du

A B S U S. s. m. Hesbe qui evolt en Egypet à la

hauteur de quelques doigts. Ses feillles reiletn-blent à celles du triolet ; & fes fleurs blanches ;

ABSYNTHE OUABSINTHE Cm. &f. Abanctium, iclos Malherbe; & felou Vange-las, toèjours matculin. On le fait plus orinnai-rement feminin. L'Academie Fr. le lait feminin. Ménage veut qu'on écrive apfinité par un p. fans doute à cause de l'étymologie. C'est une plante médecinale. Il y en a de quatte fortes : le Santonique, le marin, le vulgaire ou grand Pontique, & le petit. L'abjunte commune a une tige fort branchue, des feuilles blanches, & fore decouppées , co. . e l'artemilia ; les fleurs

fee comme une grappe de raifin. Sa racine est fort ciparpillee. Cette abjanthe vulgatee est plus en ulage dans la Medecioe. Plulicus croyent que c'est la barbotine, qu'on appelle femen fan-etum. Mais Mathiole dit que c'est une plante bien differente. Quelques-uns prétendent que l'abfinible est l'auronne femelle. Il y a une seconde espece d'abjanthe , api on appelle petre alayne , qui est semblable à la pente auronne , étane toute entaffée de petite graine fort amere . qu'on appelle en Latin abfinibium morimum ficme espèce, qui est semblable à l'aluine, mais qui est moins chargée de graise que l'autre. On fait du vin d'abjinche, & de l'eau d'abjinthe. L'abjinche est une herbe odiferante, très-amére, aftringente & corroborative : c'est pourquoy l'on s'en lett pour fortifier les viscères affoiblis: Ourre son amercume , la nicrosité dont elle participe, fait qu'elle purge la manière bilieufe con-

même en l'appliquant exteriourement. L'alfin-

PrimaLayout

lus feur est de s'en tenir aux décisions de l'E-

tion. Absolutorius. Il a une sentence absolutoire ABSORBANS, Terme de Medecine. Medicamina ad absumendum nata. Ce sont des médicamens qui ont une puillante vertu pour confamer l'humeur amollie, & attenuée fans la dif-foudre. On s'en fert ordinairement pour les tu-meurs, après avoir employé les émolliens & l'eau marine, la faumure, &c. S. Evr.

famer, emposier. Abfamere. Les eaux abfor lens presque toute la lumière qu'elles reçoives lu soleil. Ro H. Il est peu en utage au propre fi ce n'est en parlant des animaux voraces : ma l (e dic au figure , & il emporte d'ordinaire u nauvais fens. Les droits de la femme ont *abforb* xous les biens du mari, Les frais d'un (cellé abjer tent les plus clairs deniers de cette fuccession Ce goinfre a abjorbé tout son patrimoine. t eft abforbie dans les voutes. ABLAN

C'est-à-dire qu'elle s'y perd. La queition de l'infaillibilité de l'Eglise absorbe toutes les aures controverses. CLAUD. Ce mot vicos du Latin Morbeo, lignifiant le même.

fation de la peine d'un crime, abiologre, Abico are un j'énitent, lui remettre ses pechez dans le tribural de la confession. Absordre un accuse Dans le doute il est plus expédient d'absordre u eriminel, que de condamner un innocent. Count On l'a abjous à par & à plein. On dit aussi dans le discours familier, en parlant d'un dessiont : que Dieu abjalve, c'est-à-dire, à qui Dieu fasse

ssous, ourz. adj. Affranchi ou déchargé

de crime. Abfolusus. 1 25 0 u s , le dit ausli en marière civile. Un deffenueur conciue coujours a etre renvoye quitte et abjous de la demande qu'on lui a faite. A BSOUTE, f. f. Abiolution publique & fo-

Limnelle qui se donne au peuple. Dies Abjointionis. L'Evêque en fair la cérémonie le Jeudi

BS LEME. I.m. Terme dogmatique. Qu ne boit point de vin. Abstemus. On s'en fen dans la Théologie, pour parlet de ceux que dans la Communion ne peuvent point boire de

remme de vin, de fabines, &c. La dièce & l'ab-

nstinine, fignific quelquefois une fimpl desfence ou privation de manger de la chait affaitonnée de devotion, & accompagnée de l priere, est un des moyeus les plus efficaces pou avancer nôtre sanctimeation. Boss. L'Egit ordonne simplement l'abstimence le jour de lain Mare, & non paste jeune. Les Mécredis ton des jours d'abjemence chez plusieurs Religieux Les devots sont aussi des abjemences , & des ma-

BSIINENI, ENTE 30J. 1 cmperant l'égard du boire & du manger. Sobrius. Le peuples du Midi sone plus abglineus que ceux di

& de Chiturgie. Purger., nettoyer une playe

normaticus, finefficus. Medicament, purga-

BSTRACT, ACTI. Terme de Philosophie, Il est un peu barbare en François. Ce qu'on ded'en avoir une connoillance simple, & par luimeme. Species abstracta per meusem ab its rebus quibus adbarer. La quantité est un terme abrad, quand on la contidére en elle-même; & fans cire attachée à aucun corps, quoy qu'elle ne puitfe fublifier naturellement lans lui , ni lui fanselle. Lablancheur est un terme abstract . quand on la confidere détachée d'un fujet. De la connoillance des abstracts on parvient à celle des

ABSTRACTION. Lt. Cest une action de l'elprit, par laquelle on confidère quelque par-tie d'un tout, tans faire attention aux autres : ou un détachement qui le fait par la pentée de tous les accidens ou circonfrances qui peuvent ac-compagner un être , pour le confidérer mieux abil rabentes. On confidere par abiliraction, lors que dans un mobile on confidére le mouvement lans faire attention au corps mû. Ce sont les Mathématiciens qui confiderant la quantire fans matière, supposent dans leur empire d'abstrac-tion des indivisibles fans parties: mais il n'est pas permis aux Phyficiens de faire ces fortes d'ab-

lotophie. Ce qui choque le jers commun, qu'est impertirent, incroyable, impossible Abymdue. Proposition algurate. qualit on suppose une choie ab arde, on en tire mille contequences abfurdes. Il prouve une choie abfurde par une

ABU

choic plus . h a de. BSURDEMENT, adv. D'une marière ab-

BSURDITE, f. f. Qui contient quelque choic d'abiurde. Abjurde accium aut factun Il s'enfuivroit de grandes ab irrairez d'une telle fupposition. La plus grande des abjurantez est la contradiction. Quelle foi peut-on ajouter à des gens qui proposent létioulement d'aussi grandes

B S U S. L. m. Herbe dus erott en Egypte a ta hauteur de quelques doigts. Ses feuilles resfettsblent à ceiles du trioler ; & ses fleurs blanches ; & d'un jaune pale, produitent une femence noi-re, renfermée dans de petites cellules.

Abfintitium, icion Malherbe; & icion Vangelas, toù jours mateulin. On le fait plus ordinai-rement féminin. L'Academie Fr. le fait féminin Ménage veut qu'on écrive apfinihe par un p. plante médecinale. Il y en a de quatre fortes : le Santonique, le mariu, le vulgaire ou grand Pontique, & le petit. L'absparire commune a une tige fort branchuë, des feuilles blanches, & fore decouppées , co: a le l'artemilia ; les fleurs dorèes & petites, Sagraine est ronde, & dispo-tee comme une grappe de santo. Sa racine est fort espatpillée. Certe absorbe vulgaire est plus en utage dans la Modocine. Plutiours croyene que c'est la barbotine , qu'on appelle semen s'an-tum. Mais Mathiole dit que c'est une plante tien differente. Quelques-uns prétendent que l'abjenthe est l'auronne femelle. Il y a une leconde espece d'absante , 'qu'on appelle pettre alayne , qui est semblable à la pettre auronne , tant toute entaillee de petite graine fort a

u'on appelle en Latin abfinibine

l'on s'en lett pour fortifier les

u feriphisum. L'abjinche Santon itme elpéce, qui est semblable qui est moins chargée de grains ait du vin d'abunche, & de L'ablinable elt une herbe odific re, aftrionesse & corroborativ





abus de leur faux Prophete. En Arithmetique,

d'abus our été introduits aurant pour s'oppoter aux entreprites de la Juriidiction Eccleha-

ltique lur la Jurisdiction temporelle, qu'à celies de Kome jur les libertez de l'Egille Gai-

licane, Il est certain en effet que l'entreprise des

Eveques alla fi loin, qu'ils se rendirent les mai-

rres de routes les affaires civiles fous des pre-

textes de pieté , & qu'ils depouillerent pref-

qu'entierement la Jurisdiction seculière. On ne

peut point determiner rous les cas où l'on peut

appeller comme d'abds, parce qu'on ne peut pas

limiter routes les contraventions dont les Eccle-

maitiques tont capables pour relever leur auto-

INC. BOUCHEL. Labus ne le couvre point

par queique lentence , par quelque policilion ,



Expérimentation des outils

Tesseract

Librairie qui permet de faire de l'extraction de texte et de l'OCR. Pour l'utiliser, nous sommes passés par la bibliothèque pytesseract.

ABU. ne Por r-R. Co Ministre a reforme les abus A B v s z R, fignific encore interpreter mal la pensed des Finances e ce Préfident les abus de la Jude queleun, & y donner un mauvais lens. Vous abulez de quelques paroles ambiguës qui lont A Bus, lignific austi, Mauvais usage d'une thodans les lettres. Pasc. A B U S E R, fignifie ausli, Tromper, léduire. Falle. On commer bien de l'abus dans la distribution des aumônes. Les abus qu'on fait de l'Ecrilere, decipere. Les taux Prophetes, les Chariarure ne naissent pas de la secture innocente du tans abusent les peuples. Les meilleurs Arithmepeuple. GOMBER. ticiens s'abujent queiquerois en ieur caicui. A B u s , fignifie aussi , Erreur , tromperie. Error. Nôtre amour propre nous abu/e , nous tait luivic nos passions, qui nous abujent, qui nous Si vous croyez que cela loit , c'est un abus. Les Mahometans vivent dans l'abus, ils luivent les trompent. J'ai pris cet Auteur pour un autre ,

A BUSER, fignifie plus particulierement . Suquand la preuve ne le trouve pas bonne, on borner une femme, corrompre une hile, lui arconnoît qu'il y a de l'abus dans le calcul. Appel comme d'abus. In abufu dicendi juris ad racher les dernières faveurs. Vitiare. Il faut superius tribunal provocatio. C'est un appel être bien malhonnêre pour abuser de la femme qu'on interjette au Parlement, des lentences des de lon ami ; pour sbu/er de la fille de lon hô-Juges Ecclefiaftiques , quand ils entreprennent re. Etoit-il juste d'emprunter mon nom , & ma fur la putitance fecunere; quand us jugent des reliemblance pour abufer de ma Maitrefie choics qui ne lont point de leur jurisdiction , A BLANC. On s'en fert auffi dans un cas plus odieux : On dir que Neron avoit abusé plusieurs ou quand its jugent contre les baints Canons & la Discipline de l'Eglife. Les appels comme

tois de Britannicus. A BUANC

je me tuis abufe cu le citant.

trompe. Deceptor . veterator Manomet a Etc un grand abufeur de peuples.

A BUSIF. IV E. adi. Où il Vade l'abus. Errori obnoxius Une union de Benefice lans caule véritable & importante est abusive. Un jugement d'Official contre un laïque . & pour caule profane, est abusif. En termes de Grammaire, prendre un mot dans un lens abulit, c'est le placer mal; c'est en faire une mauvaise application; c'est le prendre improprement.

ABUSIVEMENT, adv. D'une manière abufive. Per abufionem. La Cour en infirmant les lentences des Juges de l'Eglife, prononce;

d'un petit Ecu qui est so milieu d'un grand qu'il est mis en stime. Et tout autant de fois qu'on commence par toute autre figure que par celle du milieu, on dit que celle qui est qui mèlieu eft en abime, comme li on vouloit dire . que les aurres grandes pièces étant élevées en relief , celle-là paroît petite , & comme cachée & abîmée. Il porte trois befans d'or avec une Reur de lis en abime. Ainfi ce terme ne fignihe pas implement le milien de l'Ecu car il eft rélatif & Inppole d'autres pieces, au milieu delquelles une plus petite eit abimee, BIME, eit aufli un vaitleau tait en prilme triangulaire renverié, qui terr aux Chandeliers à fondre leur fuif, & à faire leur chandelle , en trempant plufieurs tois leur mêche. BYSMER, v. act. Jetter dans un abime . tomber, le perdre , le noyer. In gurgitem demergere. Les Ouragans abiment les vaisseaux. Ce terrein s'est abime, il y avoit deslous une carriere. Il est quelquerois neutre: Cette ville abimera un jour à caule des abominations ou BIMER, le dir figurément en Morale, pour dire perdre ruiner entierement, Evertere, pelum dare. Les gros interets our abime ce Marchand. Ce chicaneur a abimé la partie ; il l'a ruinée de fond en comble. Il a abimé cet homme-la. Il le die plus ordinairement avec le pronom perionnel , & plus au figure qu'au propre En ce cas il marque un grand excés. C'est un volupteux qui s'abime dans les plaifirs. ACAD.FR. C'est-à-dire qui cit entièrement occu-

pe,& qui s'y abandonne fans aucune referve.

On dit, Il eft abime dans la douleur. PORT-R.

parcequ'il en est rout rempli & rout pénétré.

OCR de Tesseract

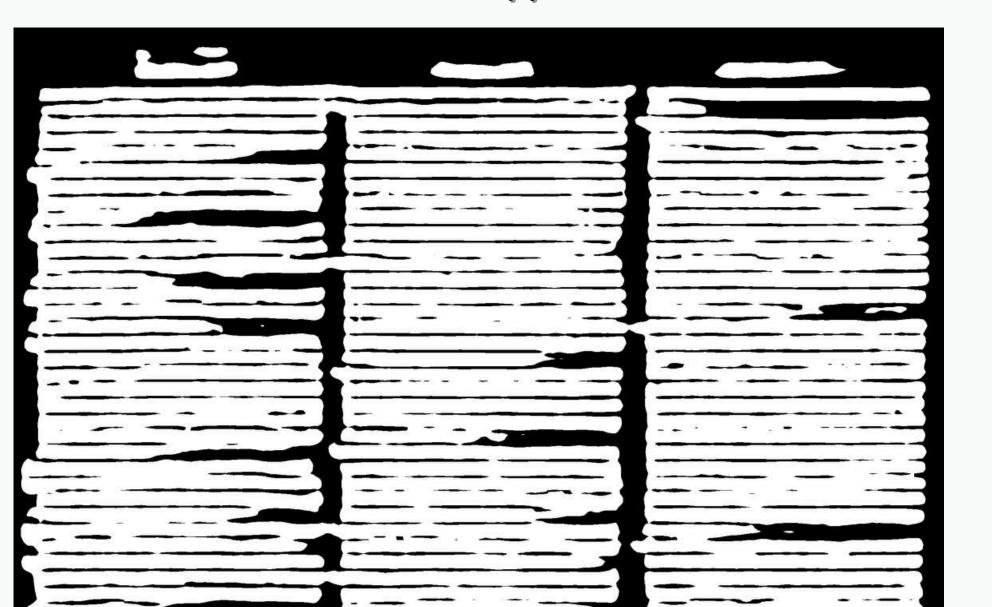
toutes les contraventions dont les Ecclep??,& qui???s???pbandonne fans





Expérimentation des outils

Résultat Laypa



Résultat surya

À.

AB. ABA.

ABA.

che: être bien à cheval. 1. La posture & legel- | Il seroit difficile de déterminer tous les differens | ABAISSER, signific aussi, Diminuër le prix. Mi-: A genoux, à bras ouverts. 3. La distance : A vingt lieues de là. 4. La qualité : De l'or à tant de carats. 5. Le prix: A dix écus. 6. La quantité: L'eau est à la hauteur d'une toile. 7. La maniere: Il est habillé à l'Espagnole. Il faut dire à coups de trait, à coups de canon: & non pas à coups de traits, & à coups de canons. MENAG. 8. La fin : Les fraudes à bonne intention ne manquent point d'approbateurs parmi les devots. Po RT-R.

A figuific successivement : Pas à pas. Il se sent mourir peu à peu. Il fignifie avec : Je l'abandonne à regret. Les douleurs à grand bruit font d'ordinaire suspectes d'affectation. M. Scub. Ce poste a été emporté à la pointe de l'épée. Peindre à l'huile,

A est plus elegant que par dans certaines phrases. Il ne faut point se laisser prendre à l'apparence, mi à l'éclat trompeur des grandeurs humaines. Flech. Ne vous laissez pas conduire à vos passions. A signific selon : A mon avis.

A Cette lettre s'employe aussi fort souvent pour marquer ce que l'on possede. C'est un homme

à caroffe, à équipage.

A se met quelquesois absolument devant l'infinitif de quelques verbes , sans être précedé d'aucun nom qui soit ou exprimé , ou sous-entendu , & alors il se peut resoudre par le gerondif. A voir les airs dedaigneux. A tout prendre, l'ailemblage de ses traits, qui sont beaux en detail, ne fait point une belle personne. Font en C'est comme si l'on disoit, en prenant tous ses traits ensemble. Passer tranquillement la nuit à bien dormir, & le jour à rien faire. Bott. Il y a aussi des occasions où il se peur resoudre par quand, ou lorsque. A ne prévoir rien on est surpris, & a prevoir trop on est miserable. S. Ev R. A raconter les maux souvent on les soulage. Co an. Il fe mer auffi devane l'infinitif de ou

ulages de la préposition ou de la particule à. On les remarquera dans la fuite, il s'en presentera des exemples presqu'à toutes les pages.

A. A. A. Les Chymittes se servent de ce signe pour fignifier , Amalgamer , Amalgamation , & Amalgame. Voyet Amalgamer.

A B A.

A B. cinquiéme mois des Hebreux, qui répond à nôtre mois de Juillet.

A B. en langue l'yriaque, le dernier mois de l'Eté.

ABA, ou ANBA, Pere, titre que les Eglises Syriaques, Cophtes & Ethiopiennes donnent à

leurs Evèques.

BACO. lubit. malc. Abacus. Ce mot le trouve wans Rouillard pour fignifier l'Arithmetique. Les Italiens disent aussi abaco pour exprimer la meme chose. C'étoit une petite table polie, sur laquelle les Anciens traçoient des figures, ou des nombres. Elle servoir à apprendre les principes de l'Arithmetique. Ils l'appelloient Table de Pythagore.

BADA. f.m. Animal farouche du païs de Benguela, dans la baile Ethiopie. Il reflemble à un cheval par la tête, & par le crin. Il est un peu moins grand. Sa queuë est pareille à celle d'un bœuf, excepte qu'elle est moins longue. Ses pieds sont fendus comme ceux du cert, & plus gros. Il a deux cornes , l'une sur le front, & l'autre fur la nuque. Les Negres tuent ces animaux à coups de fléche, pour en prendre la corne, dont ils font un remede.

BADIR. Terme de Mythologie. C'est le nom d'une pierre que Saturne dévora. Car foit parce que son frere Titanus ne lui avoit cedé l'empire du monde, qu'à condition qu'il n'éleveroit point d'enfant male; soit parce que les destinées portoient qu'il seroit un jour détrône par un de enfine "Il les faifais sons perir Enfin Cybele

nucre. Le bon ordre de la police a fait abaiffer le prix du blé; c'est-à-dire, qu'il est diminué. Ce mot en ce sens n'est pas du bel usage ; il faut dire rabaiffer. Voyez RABAISSER.

On s'en sert figurément dans le même sens. L'envie abaiffe par ses discours les vertus qu'elle peut imiter. S. Ev R. Abaiffer la Majefte du Prince. L'usage , comme la fortune , chacun dans leur jurildiction éleve ou abaiffe qui bon lui femble. V A u G. Les grands noms abaiffent, au lieu d'élever, ceux qui ne savent

pas les soutenir, Rochef. BAISSER, fignifie aussi en Morale, Ravaler l'orgueil de quelqu'un, le mortifier. Abjicere. Reprimere, Contundere. Les Romains se vantoient d'abaiffer les superbes, & de pardonner aux humbles. S.EvR. Abaiffer l'orgueil de Carthage. V A u G. Il faut abaiffer les esprits hautains. S. Eva. La crainte trouble & abaiffe l'ef-prit M. Scop. C'est-à-dire qu'elle le relache &

En termes de Fauconnerie on dit, Abaisser l'oiseau, lors qu'ayant trop d'embonpoint, on lui ôte quelque chose de son pât ordinaire, pour le mettre en état de bien voler.

BAISSIR, en termes de Jardinage, fignifie Couper une branche prés du tronc.

BAISSER, fe dit auffi avec le pronom perfonnel , & fignifie alors , s'humilier , fe foumettre , le ravaler. Abjicere fe. Il faut s'abaiffer devant la Majesté divine. S'abaisser à des choles indignes. S'abaiffer jusqu'aux plus lâches complaitances. L'humilité n'est souvent e tifice de l'orgueil, quine s'abaiffe que lever. ROCHEF. On le dit encor pect d'une personne éminente en dig

qu'elle semble rabatre de sa grander

cendant jusqu'à des personnes sort

Le Prince s'elt abaille infou'à moi







Expérimentation des outils

Tableau Récapitulatif :

	Documentation	Type-Input	Type-Output	Open source	Fine -Tunning
Layout Parser	Exhaustive	Image, PDF	JSON ,	Oui	Oui
			CSV		
Laypa	Peux exhaustive	Image	XML	Oui	Possible mais difficile
					de constituer le
					dataset
					d'entrainement
Grobid	Exhaustive	PDF	JSON,	Oui	Possible mais difficile
			XML		de constituer le
					dataset
					d'entrainement
Tesseract	Exhaustive	Image, PDF	Txt,	Oui	Non
			HOCR		
Surya	Presque	Image, PDF	JSON ,	Oui	Non
	inexistant		HOCR		



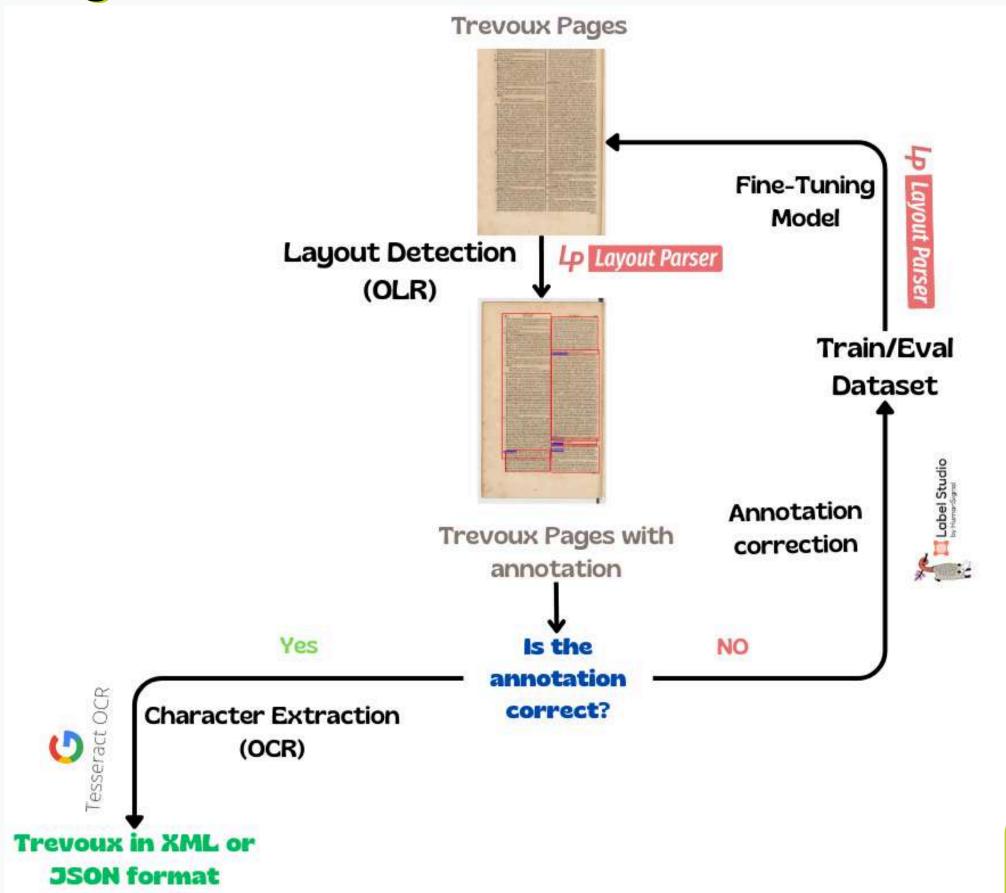


Proposition de méthodologie.

Fine-Tuner le modèle primaLayout basé sur Detectron 2, qui fonctionne assez bien sur les pages du Trévoux.

Pour ce faire, nous avons défini certaines étapes, qui sont les suivantes :

- Etape 1 : Préparation des données
- Etape 2 : Génération des annotations
- Etape 3 : Correction des annotations
- Etape 4 : Fine Tuning du model
- Etape 5 : Évaluation du nouveau model et OCR



Institut De La

communication



Résultats obtenus

Avant

AUX-AUZ-AXE-AXI-AXU-AYA-AYE-AYN-AZE-AZI-AZO-AZU.

parceque le plant est venu d'Auvergne. Leur couleur les fait appeller ailleuts Morillon, & Pineau en Auvergne, D'un Auvernat fumeux m'apporte un rouge bord. Bo 1 L. Il y a un Auvernat gris d'Orleans , qu'on appelle ailleurs Malvoifie, qui est un raisin gris fort su-cré, & le plus sondant de tous les raisins.

AUX.

UXILIAIRE. adj. m. & f. Qui vient fecours. Auxiliaris. Un prince doit plus le fie à ses soldats, qu'à ses troupes auxiliaires. Ou tre les raisons principales, on se sett quelque fois heureusement de moyens auxiliaires, o subsidiaires. Chez vous les injures sont les trou pes auxiliaires de la raison. S. E v R.

intermes de Grammaire, on appelle Verbes au xiliaires, ceux qui servent à conjuguer les au tres, comme font les verbes Etre, & Avair tant en François qu'en Italien & en Espagno Ils font communs à toutes les langues vulgaire de l'Europe.

AUZUBA. i. m. Arbre ipacieux, qui cro dans l'Isle nommée Hispaniole. Il porte un frui

AXE.

AXE. f. m. Axis. Terme de Géometrie & d'Af rronomie. C'est la ligne qui passe par le centre d'une sphere, ou d'un globe; comme la ligne qui traverse le globe de la terre. Autour de ce que one l'on sent de la douleur. Hen. Ay, ay axe tourne toute la machine du monde,& fait ioi mouvement journalier. Les deux extremite aboutifient à deux points qu'on nomme poles L'axe du monde, va d'un pole à l'autre et passant par le centre. L'axe de l'Equateur es immobile ; l'axe de l'Horizon est variable , 8 mobile. L'axe du Zodiaque en traversant la ter re, ne le termine pas aux poles du monde; le poles du Zodiaque font éloigez des poles du mon de de 23 degrez & demi. En Géometrie on dit l'a xe determine; l'axe indetermine; l'axe conjugué I'aze iccond d'une hyperbole. Voyez HYPIR-BOLE. Le grand se de l'ellipse est la ligne droi te qui en reprélente la longueur ; & le petit axe la ligne qui en représente la largeur. On obser ve dans le ciel un mouvement l'ort lent , qu'or

appelle , inclination de l'axe, on le dit auti des rouës, cônes, cylindres, & au

uses ani le menvent en rord : & on l'an-

A X U.

XUNGE, ou AXONGE. f. f. C'eft une espece de graisse la plus molle & la plus hun i-de du corps des animaux, qui s'appelle autre-ment de l'oing. Axungia Elle est differente du lard, qui est une graisse ferme; & du inf, qui est une graisse seche. Les Latins font la me ne distinction de la graisse en pinguedo, qui st l'axunge, lardum, & sevum. On l'appelie ai ssi en Latin axungia, qu'on dit avoir été fait 16 aze rotarum qua unguntur. On fe fert en M:decine de l'axonge d'oye, de canard, de vipec, & de plusieurs autres , même de celle de l'houme, qu'on estime beaucoup pour resoudre, & pour appaiser les douleurs.

XONGE DE VERRE, qu'on appelle austi fiè ou fel de verre. C'est une écume leparée de desfus

AYA

YANT. Participe du verbe Avoir , qu'on rencontre presque par tout dans les Auteurs , qui s'exprime en latin par les adverbes. Cum , Postquam, Postenguam, Ayant die cela, je m'en allai. Ayant fait beaucoup de plaintes , il se re-tira. Ayant été dangereusement blessé, il sut emporté par des foldats. Pour dire, Après avoir dit , Après avoir fait , Après avoir été bleffé.

vous me faites mal.

YEUL, sul s. f. m. & f. Pere, ou mere ceix qui ont des enfans, à l'égard desquels le nomme auffi Grand-pere, ou Grand-men Aus. Chaque enfant a un ayent paternel, ur ayeul maternel. Il fait ayeux au plurier , 5 non pas ayeuls. CORN.

Ce long amas d'ayeux que vous diffam

Sont autant de temoins qui parlent cons'e ... ZOT. f. m. Terme de Chymie. C'est aint

Quelque rang où jadis soient montez vos

AZEROLIER. f. m. Arbre fauvage, épineux, & de moyenne hauteur. Ses fenilles tone decoupées comme celles du perfil. Ses fleurs font blanches & entaflées en grapes. Il porte des fruits aigrets & fecs qu'on nomme azerolles. & qui sont rouges & gros comme des cerises. Ils sont assez agreables au goût étant meurs. C'est une espece de néssier. En Latin mespilus apit folio laciniato, ou mespilus Aronia. On le greffe sur l'épine blanche, ou sur le sauvageon de poir et, & sur le cognacier. Il y en a un qui viene du Conselus des foires sont de Conselus des foires sont de comme de conselus des sont des foires sont de conselus des sont des foires sont de conselus vient du Canada, dont les épines sont treslongues, & les feuilles très-grandes. Il yenz aufli un blanc qui vient de Florence, qu'on ne trouve qu'à Versailles , & qui ne differe de l'autre que par la couleur de fon fruit.

ALI

ZIMUI. I. m. Termed Attronomie. Ceft un grand cercle vertical qui passe par le zenith & le nadir, & qui couppe l'horison à angles droits. Verticalis circulus horiZontem ad angulos rectos interfecans. Or comme l'horifon eft divise pars 60 degrez, il donne lieu à decrire 360. azimurs. Ce mot eft purement Arabe. Ces cereles sont les memes que les rumbs des Mariniers marquez fur la Carre. On commence à les compter depuis le point du vrai Orient ou de l'Orient Equinoctial, & on continue en allanvers le Midy jusqu'à 360. C'est dans ces cercles qu'on prend la hauteur des astres à toutes les

AZIMUTAL, ALE, adj. Terme d'Aftre nomie. Il ne se dit d'ordinaire qu'au malculin. Il fignifie qui représente, ou qui mesure les azimuts. Quod verticales circulos exhibet. Un cercle azimutal; c'est celui qu'on s'imagine être mene du point vertical fur l'horiton à angles droits. On dit aussi, Cadran az imutal; & c'est celui dont le stile est à angles droits sur le

que les Chymistes appellent la matiere premier des méraux.

Leur gloire de si loin n'éblouit point mes l'Isle de Madagascar. Il croît sur un petit at brisseau comme les groseilles.

L'amour de vos ayeux passe en vous po er / ZOUFA. s. f. Bete du Royaume de Casubi Both On en trouve autu a ser, On en trouve aush à Fez , & à Maroc. Ce

Apres

AUX-AUZ-AXE-AXI-AXU-AYA-AYE-AYN-AZE-AZI-AZO-AZU.

A X U.

parceque le plant est venu d'Auvergne. Leur

couleur les fait appeller ailleurs Morillon , &

Auvernat gris d'Orleans, qu'on appelle ail-leurs Malvoisse, qui est un raisin gris fort su-cré, & le plus fondant de tous les raisins.

AUX.

TXILIAIRE adj. m. & f. Qui vient at

fecours. Auxiliaris. Un prince doit plus fe fier

à ses soldats, qu'à ses troupes auxiliaires. Ou-tre les raisons principales, on se sers quelque-

fois heureusement de moyens auxiliaires, ou

subsidiaires. Chez vous les injures sont les trou-

En termes de Grammaire, on appelle Verbes au-

xiliaires, ceux qui servent à conjuguer les au-

tres, comme font les verbes Etre, & Avoir,

tant en François qu'en Italien & en Espagnol.

Ils font communs à toutes les langues vulgaires

AUL.

ZUBAIf. m. Arbre Ipacieux, qui croi

A X E.

E. 1 m. Axis. Terme de Géometrie & d'Al-

rronomie. C'est la ligne qui passe par le centre d'une sphere, ou d'un globe; comme la ligne qui traverse le globe de la terre. Autour de cet

axe tourne toute la machine du monde, & fait son

mouvement journalier. Les deux extremitez

aboutissent à deux points qu'on nomme poles. L'axe du monde, va d'un pole à l'autre en

passant par le centre. L'axe de l'Equateur est

immobile ; l'axe de l'Horizon est variable , &

mobile. L'axe du Zodiaque en traversant la ter-

ze, ne le termine pas aux poles du monde; les

poles du Zodiaque sont éloigez des poles du mon-de de 23 degrez & demi. En Géometrie on dit l'a-

xe determine; l'axe indetermine; l'axe conjugué;

Taxe tecond d'une hyperbole. Voyez HYPER-

BOLE. Le grand axe de l'ellipse est la ligne droi-

te qui en représente la longueur ; & le petit axe,

la ligne qui en représente la largeur. On observe dans le ciel un mouvement fort lent , qu'on

appelle, inclination de l'axe, On le dit aush des rouës, cônes, cylindres, & au-

eres figures qui le meuvent en rond ; & on l'ap-

dans l'Iste nommée Hispaniole. Il porte un frui

pes auxiliaires de la raison. S. E v R.

Pineau en Auvergne. D'un Auvernat fumeux stre UNGE ou utre ONGE. t. C'est une m'apporte un rouge bord. Bo 1 L. Il y a un espece de graisse la plus molle & la plus humide du corps des animaux, qui s'appelle autre-ment de l'oing. Axungia Elle est disserente du lard, qui est une graisse ferme; & du suif, qui est une graisse seche. Les Latins font la même distinction de la graisse en pinguedo, qui est l'axunge, lardum, & sevum. On l'appelle aussi en Latin axungia , qu'on dit avoir été fait ab axe rotarum que unguntur. On le fert en Medecine de l'axonge d'oye, de canard, de vipere, & de plufieurs autres , même de celle de l'homme, qu'on estime beaucoup pour resoudre, & pour appaiser les douleurs.

XONGE DE VERRE, qu'on appelle aufli fiel, ou sel de verre. C'est une écume leparée de dessus la matiere du verre, avant qu'elle se vitrifie.

AYA.

ANT. articipe du verbe Avoir, qu'on rencontre preique par tout dans les Auteurs, qui s'exprime en latin par les adverbes. Cum , Poftquam. Posteaquam. Ayant de cela, je m'en allai. Ayant fait beaucoup de plaintes, il se retira. Ayant été dangereusement blesse, il sut emporté par des foldats. Pour dire, Après avoir dit, Après avoir fait, Après avoir été blessé.

A Y-E.

I'I', AYE. Interjection, Exclamation qui marque que l'on fent de la douleur. Heu. Ay , mye

HIM EUL Bul s. l. m. & f. Pere, ou mere de ceux qui ont des enfans, à l'égard desquels on les nomme aufli Grand-pere, ou Grand-mere. Avus. Chaque enfant a un ayeul paternel, & un ayeul maternel. Il fait ayeux au plurier, & non pas ayeuls. CORN.

Ce long umas d'ayeux que vous diffame?

Sont autant de temoins qui parlent contre Quelque rang où jadis soient montez vos

Leur gloire de si loin n'éblouit point mes

L'amour de vos ayeux passe en vous pour BoiL

AZEROLE. f. f. Fruit d'Azerolier, EROLIER I. m. Arbre sauvage, épineux, & de moyenne hauteur. Ses femilles font decoupées comme celles du perfil. Ses fleurs font blanches & entaffées en grapes. Il porte des fruits aigrets & fecs qu'on nomme azerolles. & qui sont rouges & gros comme des cerifos. Ils sont assez agreables au goût étant meurs. C'est une espece de néssier. En Latin mespilus apii folio laciniato, ou mespilus Aronia. On le greffe sur l'epine blanche, ou sur le sauvageon de poirier, & sur le cognacier. Il y en a un qui vient du Canada, dont les épines sont très-longues, & les feuilles très-grandes. Il yena aufli un blane qui vient de Florence, qu'on ne trouve qu'à Versailles, & qui ne dissere de l'au-tre que par la couleur de son fruit.

AZI.

IMUT. I. m. Termed Aitronomie. C'eft un grand cercle vertical qui passe par le zenith & le nadir, & qui couppe l'horison à angles droits. Verticalis circulus horizontem ad angulos rectos intersecans. Or comme l'horison est divife pars 60 degrez, il donne lieu à decrire 360. azimurs. Ce mot est purement Arabe. Ces cercles sont les mêmes que les rumbs des Mariniers marquez fur la Caite. On commence à les compter depuis le point du vrai Orient ou de l'Orient Equinochial, & on continue en allanvers le Midy jusqu'à 360. C'est dans ces cereles qu'on prend la hauteur des astres à toutes les

nomie. Il ne se dit d'ordinaire qu'au masculin. Il fignifie qui représente, ou qui mesure les azimuts. Quod verticales circulos exhibes. Un cercle azimutal; c'est celui qu'on s'imagine être mene du point vertical sur l'horison à angles droits. On dit aussi, Cadran azimutal; & c'est celui dont le stile est à angles droits sur le plan de l'horison.

AZO.

OT. f. m. Terme de Chymie. C'est ainsi que les Chymistes appellent la mariere première des métaux.

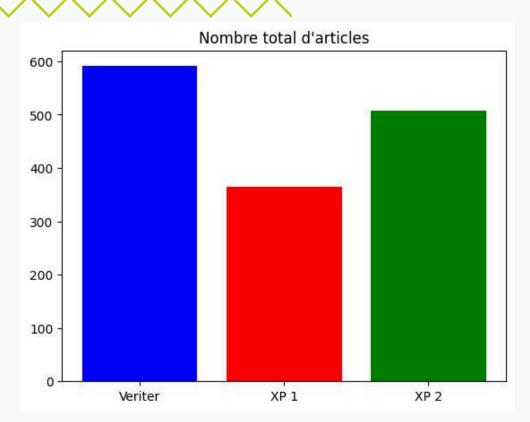
l'Isle de Madagascar. Il croît sur un petit arbriffeau comme les groseilles.

On en trouve aussi à Fez, & à Maroc. Ces Se pare qui voudra du nom de ses aveux? animaux deterrent les morts. & les dé-





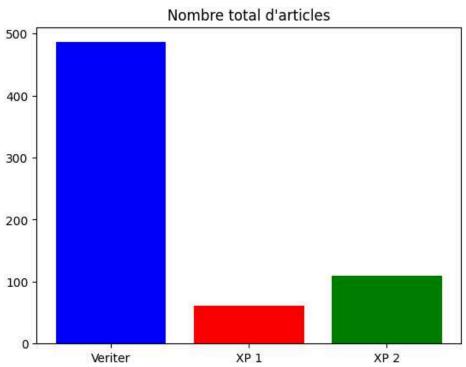
Évaluation des résultats



Interprétation:

le modèle XP2 est plus performant que le modèle XP1 en termes de détection d'articles.

En effet, le modèle XP1 a détecté 364 articles sur 591 possibles, soit un ratio de 0,615, alors que XP2 a détecté 508 articles sur 591, soit un ratio de 0,859



Interprétation:

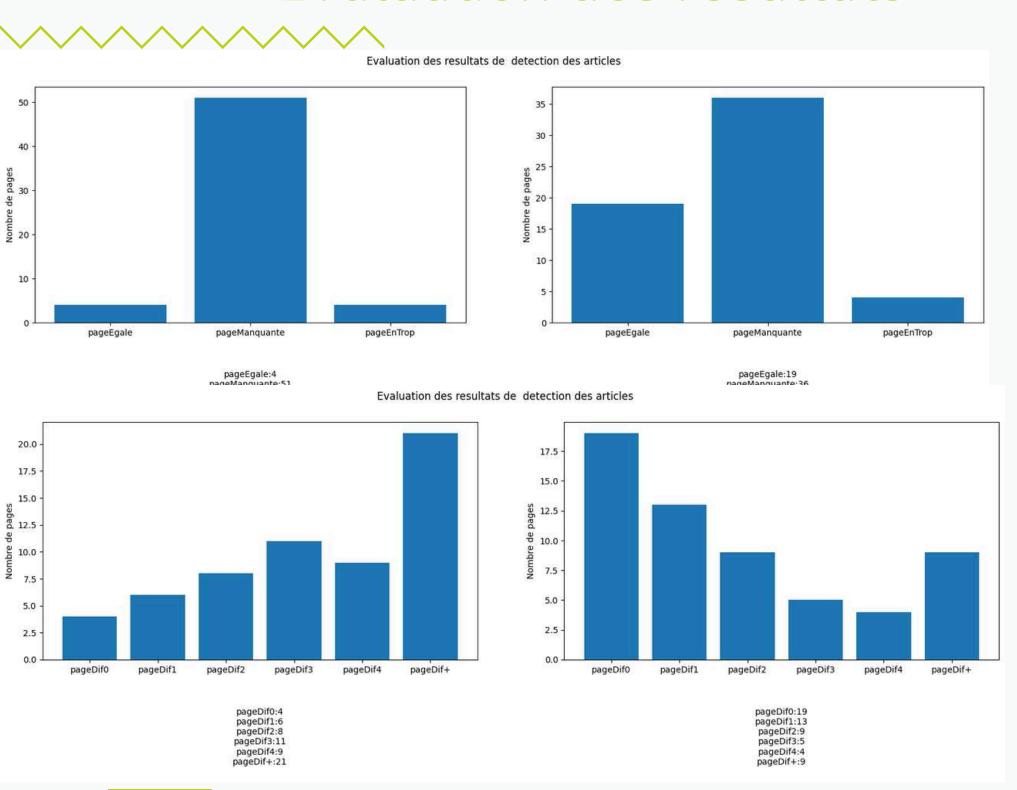
On remarque que les deux modèles sont relativement mauvais dans la détection de vedettes.

En effet, le modèle XP1 a détecté 61 vedettes sur 486 possibles, soit un ratio de 0,125, alors que XP2 a détecté 110 vedettes sur 486, soit un ratio de 0,22.





Évaluation des résultats



Interprétation:

En zoomant sur le nombre de détections manquantes et en trop par page d'article, on remarque qu'il y a dans la plupart des cas une sous-détection, sauf dans le cas de 4 pages où les 2 modèles font de la surdétection

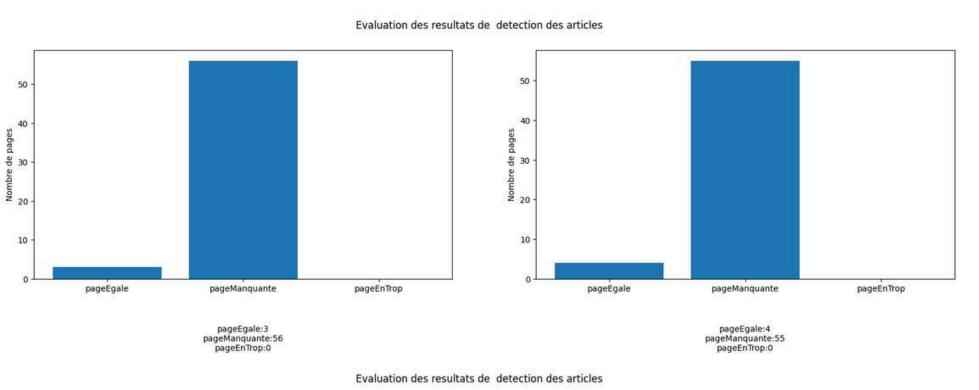
Interprétation:

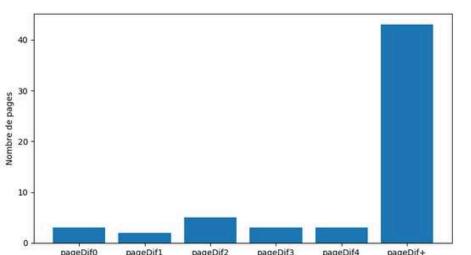
Ces graphiques montrent que le modèle XP2 (à droite) est beaucoup plus efficace dans la détection d'articles, car il fait le moins d'erreurs contrairement à XP1.

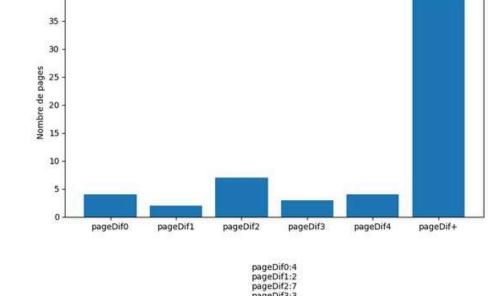




Évaluation des résultats







Interprétation:

En zoomant sur le nombre de détections manquantes et en trop par page de vedette, on remarque qu'il y a dans la plupart des cas une sous-détection donc nos 2 modèle ne sont pas assez bonne pour détecter les vedettes

Interprétation:

Dans la détection de vedettes, il n'y a pas de véritable tendance qui se dégage, car les deux modèles semblent être aussi mauvais l'un que l'autre.





Évaluation des résultats

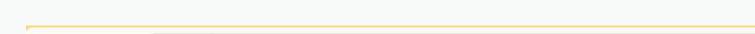


- **Prédiction avec notre modèle :** Réaliser une prédiction sur un échantillon du Trévoux.
- Extraction du texte : Utiliser l'OCR Tesseract pour extraire les textes.
- Reconstitution des prédictions : Reconstituer les prédictions pour chaque page.
- Préparation de la vérité terrain : Reconstituer l'échantillon de vérité terrain par page à l'aide des marqueurs disponibles dans nos fichiers.
- Jointure des données : Effectuer une jointure entre les prédictions et la vérité terrain en utilisant la clé des pages.
- Alignement des données : Réaliser un alignement grâce à l'algorithme RETAS.
- Affichage et comparaison des résultats : Afficher et comparer les résultats obtenus.





Évaluation des résultats



A D ≡

25donnera sera clair, ou obscur, ou obtus, bref ou long, selon les diffA@rentes 26 consonnes qui le suivront, comme on vient de l'expliquer. A s. m. C'est le nom 27de cette lettre, ou du caractà re que nous appellons a. Un grand a, un petit a,

Fichier /home/eadjaro/Images/Fine-Tuning/Evaluation/EvaluerOCR/ParPage/temp.html

28un a bien formÃC. Ce nom est du genre masculin, comme celui de toutes les 29voyelles Franã§oises. Cette lettre sert de corps à un Rebus en cette maniÃ"re : 30On range plusieurs A de suite jusqu'Ã un tombeau, & ces paroles font l'ame 31du Rebus, Amis jusqu'au tombeau. Cette lettre A ACtoit aussi chez les Anciens 32une lettre numÃ0rale qui signifioit 500. comme on le voit dans Valerius Probus. 33 Voyez sur ces prétendues lettres numérales ce qu'on en a remarqué sur la 34 lettre e. Il y a des vers anciens rapportÃOs par Baronius, qui marquent les 35 lettres significatives des nombres, dont le premier est : Possidet A numeros 36 quingentos ordine recto Quand on mettoit un titre ou une ligne droite au-dessus 37de l'A, il significit cinq mille. Les Romains l'appelloient lettre salutaire, 38 parce qu'on s'en servoit pour dacclarer innocent celui qui actus accusac. A 39 vouloit dire absolvo, je l'absous. Cette lettre a diverses significations. 40 Cependant il en faut Acviter la rencontre trop fracquente dans une maame 41période. Quelquefois cette répétition rend le discours rude & moins 42agréable. C'est quelquefois un substantif masculin. Cet A est mal formé. On 43dit par une façon de parler proverbiale : il n'a pas fait une pan<mark>se</mark> d'a, pour 44dire, il n'a pas formé une seule lettre, & figurément, il n'a fait quoi 45que ce soit. On dit aussi dans la conversation familià re : Il ne sait ni A ni 46B, pour exprimer un ignorant. Ci-dessous gît Mr l'Abbé, Qui ne savoit ni A ni 47B. MÃ@nag. C'est aussi la troisià me personne du verbe auxiliaire avoir. Il a 48 fait de l'éclat mal-à -propos. L'imagination du Poëte n'a pu vous peindre si 49belle que vous êtes. Voit. La vérité, qui a des bornes, a dit pour vous tout 50ce que le mensonge, qui n'en connoît point, a inventé pour les autres. S. Evr. 25d'un mot fa prononciation chan- gera, &@@@@ le fon qu'on lut donnera fera clair 26, ou obfcur , ou obtus, bref ou long , felon les ditf@érentes confonnes qui le 27 {uivront, comme on vient de @@expliquer. . A £, m. C@eft@le nom de cette letere 28, ou du caract@ére que nous appelions 4, Un@grand 4, un pe@nt@a, un # bien 29 form@é. Ce nom eft du genre@mafculin, comme celui de toutes les voyel- 1es 30 Fran@goites. Ceste lettre fert de corps@d@un Rebus en cette mani@ére : On range 31pluficurs 4 de funite jufqu'd un tombeau, &@@@@ ces paroles font @'ame du Rebus, 32Amis jufqu'an tonbean. Cetze lettre # @éroir aufli chez les Anciens une le@wre 33 num@érale qui {ignifioit §00, comme on le voit dans Valerius Probus. Voyez fur 34ces@pr@étendues le@twes num@érales ce qu@on@en@a@re- marqu@é @(i la le@rre@e@ @e 351y a des vers anciens rapport@és par Baroniuss , qui marquent lgs lerrres 36 fignificatives des nom- bres, dont le premier eft : Poffider A numeros 37 quingentos ordine reclo Quand on mettoit un titre ou une ligne droite au-deffus 38de @T4, il fig@mf@oir cing mille. Les Romains l'a@pfinlbi@cm lettre Salutaire, 39parce qu@on s'en fervoit pour d@éc@@@@@ innocent *"celui qui @éroit accu@lé. A 40 vouloit dire abfolve, je @Iablous. Cetre le@tre 2 diverfes fignifications. 41Cependant il en faut @évi- « ter la rencontre grop fr@équente dans une m@éme 42p@ériode. . Quel _{flxcfois cexce r@ép@@éution rend le difcours rude &@@@@ moin: 43agr@éable. = .Cclbt quickquefois un fubftantif mafculin. Cet A cft mal formé " 440n dit par nne fa@gon de parler proverbiale : il n'a pas fai@ une pan@c diur, 45 pour dire , il n'a pas form@é une feule lettre: &@@@@ figur@ément , 1l r'a 46 fair@quoi que ce foit. On dit aufli dans la converfarion famili@ére : Il ne fai@ 47 ni@4 ni B , pour expri . Aner un ignorank, S U Ci-deffous g@it Mr I Abb@é, e 48Qai@fic@[avf}t@@idA@ni@B.'M@fiNAc. i C@eft aulli la tro@@ifi&m% perfonne du verbe 49auxiliaire avoir. 11 . "~ faic de l'@éclar mal-a@-propos. L'imagination du 50Po@cte n'a pt wous peindre fi belle que: vous @&tes. Vorr. La v@érité , qu * a 51des bornes , « dit pour vous rout ce que le menfonge, qu _ n@en conno@it point,





Évaluation des résultats

L'analyse des résultats obtenus révèle que :

- Il y a une amélioration significative dans la détection des contours.
- Les contours prédits par nos modèles ne sont pas toujours parfaits, ce qui entraîne des parties de texte manquantes.
- Notre OCR Tesseract a tendance à commettre de nombreuses erreurs, notamment des confusions avec les "s" longs.
- De plus, notre OCR Tesseract a également tendance à halluciner, c'est-àdire à reconnaître des éléments inexistants.





Conclusion et perspectives

Conclusion

L'expérimentation des outils qui nous ont été proposés, tels que Tesseract, Layout Parser, Laypa, Grobid et Surya, nous a permis de découvrir un outil intéressant : Layout Parser avec Detectron2, qui semble, dans un premier temps, être le plus efficace sur Trevoux nous a inspirer une méthode.

Après l'expérimentation de cette méthodologie, nous avons constaté une nette amélioration des résultats de notre outil, et la marge de progrès semble significative au vu de la quantité de nos échantillons.

De plus, ce stage m'a permis de m'initier au monde de la recherche, mais aussi de développer différentes compétences, parmi lesquelles la prise en main rapide d'un outil grâce à sa documentation, ainsi que l'apprentissage de la bonne documentation d'un travail.





Conclusion et perspectives

Perspectives:

- Amélioration des échantillons d'entraînement: Continuer à enrichir et perfectionner la qualité et la quantité de nos échantillons d'entraînement afin d'optimiser les performances du modèle.
- Entraînement de sous-modèles: Pour améliorer nos détections, l'entraînement de plusieurs sous-modèles, à l'instar du modèle XP2, pourrait nous permettre d'obtenir des résultats plus précis et fiables.
- Exploration d'autres OCR: Explorer d'autres OCR spécifiquement adaptés aux textes en français ancien pourrait également constituer une piste prometteuse pour améliorer la précision des reconnaissances.





Mercil

Présenté par Edina ADJARO PATOUSSI 30 Août 2024

Tuteur laboratoire:

Tuteur enseignant:

Julien VELCIN

Ludovic MONCLA

Valentin Lachand-Pascal

